True Model: $Y_i = \alpha + \beta x_i + u_i$ [α, β = unknown popln parameters]

Estimated Model: $\hat{Y}_i = \hat{\alpha} + \hat{\beta} x_i$

Obj : Find $\hat{\alpha}, \hat{\beta}$ using the OLS Estimation technique to obtain the estimated model [good approx of the true model (unknown)].

OLS: $\underset{\hat{\alpha}, \hat{\beta}}{\text{Min}} \sum_{i=1}^{n} e_i^2$, $e_i = (Y_i - \hat{Y}_i)$ = error in estimation.

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = \sum (Y_i - \hat{\alpha} - \hat{\beta} x_i)^2$$

$\therefore$ For min:

$$\frac{\partial \sum e_i^2}{\partial \hat{\alpha}} = 0 \Rightarrow (-\cancel{2}) \sum (Y_i - \hat{\alpha} - \hat{\beta} x_i) = 0 .$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}} = 0 \Rightarrow (-\cancel{2}) \sum (Y_i - \hat{\alpha} - \hat{\beta} x_i) x_i = 0 .$$

$\rightarrow$ 2 eqns (normal eqns). to solve for 2 unknowns $\hat{\alpha}, \hat{\beta}$.

$$(i) \Rightarrow \sum \overbrace{(Y_i - \hat{\alpha} - \hat{\beta} x_i)}^{e_i} = 0 \Rightarrow \boxed{\sum e_i = 0}$$

$$\Rightarrow \sum Y_i - n\hat{\alpha} - \hat{\beta} \sum X_i = 0 .$$

$$\Rightarrow \bar{Y} - \hat{\alpha} - \hat{\beta} \bar{X} = 0 \Rightarrow \hat{\alpha} + (\bar{X}) \hat{\beta} = \bar{Y} \quad \text{.....} \quad NE(i)$$

$$(ii) \Rightarrow \sum (Y_i - \hat{\alpha} - \hat{\beta} x_i) x_i = 0 . \Rightarrow \boxed{\sum e_i x_i = 0}$$

$$\sum Y_i X_i - \hat{\alpha} \sum X_i - \hat{\beta} \sum X_i^2 = 0 .$$

$$\sum Y_i X_i = \hat{\alpha} \sum X_i + \hat{\beta} \sum X_i^2 \Rightarrow (\sum X_i) \hat{\alpha} + (\sum X_i^2) \hat{\beta} = \sum Y_i X_i \quad \text{...} NE(ii)$$

Start with $\hat{\beta}$ using Cramer's Rule:

$$\hat{\beta} = \frac{\begin{vmatrix} 1 & \bar{Y} \\ \sum x_i & \sum Y_i x_i \end{vmatrix}}{\begin{vmatrix} 1 & \bar{X} \\ \sum x_i & \sum x_i^2 \end{vmatrix}}$$

$$a_1 x + b_1 y = c_1 .$$
$$a_2 x + b_2 y = c_2 .$$

$$x = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}} \qquad y = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}$$

$$= \frac{(\sum x_i Y_i - \bar{Y} \sum x_i)/n}{(\sum x_i^2 \quad \bar{X} \sum x_i)/}$$

$$= \frac{\left(\sum x_i Y_i - Y \sum x_i\right)/n}{\left(\sum x_i^2 - \bar{x} \sum x_i\right)/n}$$

$$= \frac{\frac{1}{n}\sum x_i Y_i - \bar{x}\,\bar{Y}}{\frac{1}{n}\sum x_i^2 - \bar{x}^2} = \frac{\cancel{\frac{1}{n}}\sum (x_i-\bar{x})(Y_i-\bar{Y})}{\cancel{\frac{1}{n}}\sum (x_i-\bar{x})^2} = \boxed{\frac{\sum (x_i-\bar{x})(Y_i-\bar{Y})}{\sum (x_i-\bar{x})^2} = \hat{\beta}}$$

From NE(i) : $\boxed{\hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{x}}$

$$\boxed{\text{OLS Estimates: } \hat{\beta} = \frac{\sum (x_i-\bar{x})(Y_i-\bar{Y})}{\sum (x_i-\bar{x})^2}, \quad \hat{\alpha} = \bar{Y} - \hat{\beta}\,\bar{x}}$$

Properties of the OLS Estimates:-

(1) From the NE's we get: $\sum e_i = 0$, $\sum e_i x_i = 0$.

$\sum e_i = 0 \Rightarrow \bar{e} = 0$.

$$\mathrm{cov}(X, e) = \frac{1}{n}\sum (x_i-\bar{x})(e_i-\bar{e})$$

$$= \frac{1}{n}\sum (x_i-\bar{x})\,e_i \quad [\because \bar{e} = 0]$$

$$= \frac{1}{n}\underbrace{\sum x_i e_i}_{=0} - \frac{1}{n}\,\bar{x}\underbrace{\sum e_i}_{=0} = 0.$$

$\mathrm{corr}(x, e) = 0$. [Explanatory variable & error are uncorrelated].

(2) $$\sum \hat{Y}_i e_i = \sum (\hat{\alpha} + \hat{\beta} x_i)\,e_i$$

$$= \hat{\alpha}\underbrace{\sum e_i}_{=0} + \hat{\beta}\underbrace{\sum x_i e_i}_{=0} = 0.$$

$$\mathrm{cov}(\hat{Y}, e) = \frac{1}{n}\sum (\hat{Y}_i - \bar{\hat{Y}})(e_i - \bar{e})$$

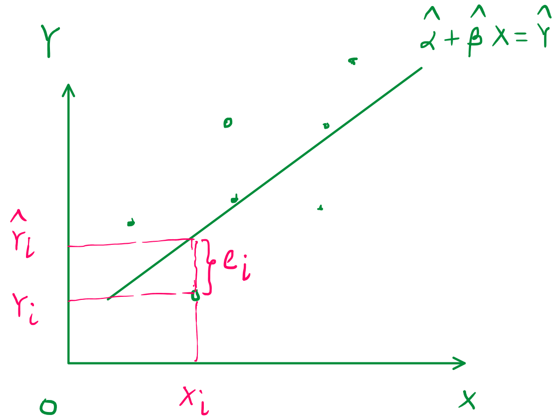
$$= \frac{1}{n}\sum (\hat{Y}_i - \bar{\hat{Y}})\,e_i$$

$$= \frac{1}{n}\underbrace{\sum \hat{Y}_i e_i}_{=0} - \bar{\hat{Y}}\cdot\frac{1}{n}\underbrace{\sum e_i}_{=0} = 0$$

$$= \frac{1}{n}\underbrace{\left(\sum r_i e_i\right)}_{0} - r \cdot \frac{1}{n}\underbrace{\left(\sum e_i\right)}_{0} = 0$$

$corr(\hat{Y}, e) = 0$ [Estimated value is uncorrelated with the error value].

$\Rightarrow$ OLS Estimated model is kind of the best-fitted model to the data.



X causes Y

True Model: $Y_i = \alpha + \beta X_i + u_i$

Estimated Model: $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$

$\Rightarrow$ understand / explain the variability in Y through variability in X.

HW

(3) Show: $\bar{Y} = \bar{\hat{Y}}$

(4) True Model: $Y_i = \alpha + \beta X_i + \gamma X_i^2 + u_i$.

Can this model be estimated using OLS? [Can we find $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$]